Molecular composition of dissolved organic matter in the Mediterranean Sea

Alba María Martínez–Pérez*[,1], Helena Osterholz[2], Mar Nieto–Cid[1], Marta Álvarez[3], Thorsten Dittmar[2], Xosé Antón Álvarez–Salgado[1]

[1] Consejo Superior de Investigaciones Científicas – Instituto de Investigacións Mariñas (CSIC–IIM), Vigo, Spain

[2] Research Group for Marine Geochemistry, Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University, Oldenburg 26129, Germany

[3] Instituto Español de Oceanografía (IEO), Centro Oceanográfico de A Coruña, A Coruña, Spain

*Corresponding author: albam@iim.csic.es

Instituto de Investigacións Mariñas

Eduardo Cabello, 6, 36208, Vigo, Pontevedra, España

Tel.: +34986231930 – Ext 364

Keywords: Dissolved organic matter, molecular composition, FT–ICR–MS, Mediterranean Sea

Abstract

The molecular composition of marine dissolved organic matter (DOM) is still poorly understood, particularly in the Mediterranean Sea. In this work, DOM from the open Mediterranean Sea and the adjacent Northeast Atlantic Ocean was isolated by solid–phase extraction (SPE–DOM) and molecularly characterized using Fourier–Transform Ion Cyclotron Resonance Mass Spectrometry (FT–ICR–MS). We assessed the gradual reworking of the SPE–DOM transported by the shallow overturning circulation of the Mediterranean Sea by following the increase in molecular weight (+20 Da), oxygenation (+5%), degradation index (Ideg +22%) and the proportional decrease of unsaturated aliphatic compounds (+34%) along the Levantine Intermediate Water (LIW). This reworked SPE–DOM that leaves the Mediterranean Sea through the Strait of Gibraltar strongly contrasts with the fresh material transported by the inflow of Atlantic water (Ideg –25%). In the deep eastern and western overturning cells the molecular composition of the deep waters varied according to their area and/or time of formation. SPE–DOM of the waters formed in the Aegean Sea during the Eastern Mediterranean Transient (EMT) was more processed than the DOM in pre–EMT waters formed in the Adriatic Sea (molecular weight and the proportion of unsaturated aliphatic compounds were increased by 5 Da and 9%, respectively). Furthermore, pre–EMT waters contain more reworked SPE–DOM (Ideg +7%) than post–EMT waters formed also in the Adriatic Sea. In summary, our study shows that the Mediterranean Sea constitutes a laboratory basin where degradation processes and diagenetic transformations of DOM can be observed on close spatial and temporal scales.

Introduction

Marine dissolved organic matter (DOM) is one of the largest and least understood reservoirs of reduced carbon on the Earth's surface (Hansell, 2002, Hedges, 1992). At 662 Pg C, DOM represents 96% of the total organic carbon in the oceans (Hansell et al., 2009). It is produced mainly in the epipelagic layer (0–150 m depth) as a result of phytoplankton photosynthesis and subsequent food web interactions (Carlson, 2002). Most of this recently produced DOM is quickly respired to CO_2. However, a small fraction of this material escapes rapid mineralization, accumulating in the upper layers for eventual export to the dark ocean (>150 m depth) by convective overturning and vertical mixing (Hansell et al., 2009).

To achieve a better understanding of the fate of DOM in the dark ocean, identifying the molecular composition and structure of this material is essential. Previous studies applying ^{1}H nuclear magnetic resonance (NMR), amino acid and neutral sugar analysis of ultrafiltered DOM (UDOM) revealed that carbohydrates are main constituents of this material at the sea surface of the Mediterranean Sea (Jones et al., 2013). This pool decreases with depth indicating DOM biodegradation. In addition, a strong correlation between amino acid concentration, apparent oxygen utilization (AOU) and picoplankton activity has been observed (Jones et al., 2013; Meador et al., 2010).

Tangential–flow ultrafiltration with a 0.5–1 kDa cut–off is able to isolate up to 30% of marine DOM (Amon and Benner, 1996; Benner et al., 1992; Benner et al., 1997). Solid–phase extraction (SPE) using styrene divinyl benzene polymer (PPL) cartridges has more recently been introduced as an efficient method for isolating more than 60% of marine DOM (Dittmar et al., 2008; Green et al., 2014). The salt–free extracts are accessible by modern, non–targeted ultrahigh–resolution analytical

techniques such as Fourier–Transform Ion Cyclotron Resonance Mass Spectrometry (FT–ICR–MS) for a comprehensive characterization. Nowadays, FT–ICR–MS is a widely used technique to distinguish thousands of molecular formulae constituting the DOM pool. Previous studies on the molecular composition of DOM by FT–ICR–MS showed molecular level differences between terrestrial and marine DOM (Koch et al., 2005), open ocean and coastal DOM (Koprivnjak et al., 2009) as well as surface and deep water DOM in the North Pacific (Medeiros et al., 2015) and North Atlantic (Hansman et al., 2015) oceans. Furthermore, the effect of degradation on the molecular composition of DOM was investigated along the eastern Atlantic and Southern Oceans combining FT–ICR–MS with radiocarbon analysis (Flerus et al., 2012; Lechtenfeld et al., 2014). Hertkorn et al. (2006) combined multidimensional NMR with FT–ICR–MS on ultrafiltered DOM, reporting carboxylic–rich alicyclic molecules (CRAM) as a likely major component of the DOM (8% of the whole DOM pool). More studies exist on the molecular characterization of open ocean DOM by FT–ICR–MS (Chen et al., 2014; Hansman et al., 2015; Hertkorn et al., 2006; Medeiros et al., 2015), but DOM composition in the enclosed Mediterranean Sea has not been studied in this detail yet.

The Mediterranean Sea is considered a concentration basin (evaporation > precipitation + runoff) characterized by low nutrient concentrations. This is due to the imbalance between the bottom outflow of nutrient–rich Mediterranean Water and the surface inflow of nutrient–poor Atlantic water at the Strait of Gibraltar (Huertas et al., 2012). High oxygen concentrations in the deep layers are a consequence of the recent formation of the Mediterranean deep waters (Cruzado, 1985). Relatively small in size, the Mediterranean Sea has been used as a test basin for general ocean circulation studies (Bergamasco and Malanotte–Rizzoli, 2010; Béthoux et al., 1998). The time scale of the

Mediterranean Sea meridional overturning circulation is about 50–80 years (Pinardi and Masetti, 2000), compared to about 350 years for the world ocean (Laruelle et al., 2009).

During the HOTMIX cruise (April–May 2014), we collected samples at selected depths along a longitudinal transect from the Levantine Sea to the Northeast Atlantic Ocean to characterize the molecular composition of solid–phase extractable DOM (SPE–DOM) through the water column *via* FT–ICR–MS. The specific objectives of our study are to 1) determine the overall molecular composition of the SPE–DOM in the Mediterranean Sea; 2) compare the molecular composition between the inflow of Atlantic surface water entering the Mediterranean Sea and the overflow of Mediterranean water into the Atlantic Ocean; and 3) explore the main drivers controlling the DOM transformations through changes in the molecular characteristics of SPE–DOM in relation to the "oceanographic model system" of the Mediterranean Sea.

Materials and methods

Study site

The Mediterranean Sea is a semi–enclosed basin opened to the Atlantic Ocean through the Strait of Gibraltar. It is constituted by two basins of similar size, western and eastern, connected *via* the Strait of Sicily. The main water masses observed in the Mediterranean Sea are the Atlantic water (AW) in the epipelagic layer, the Levantine Intermediate Water (LIW) in the mesopelagic layer and the Eastern (EMDW) and Western (WMDW) Mediterranean Deep waters in the bathypelagic layer. The Atlantic inflow enters the Strait of Gibraltar as a surface current of salinity (S) about 36.5, being slightly modified through evaporation and mixing with the outflowing Mediterranean waters, leading to the Modified Atlantic Water (MAW), which moves towards the East through a shallow and open thermohaline cell that spans the two basins and leads to the

formation of intermediate waters in the eastern basin (Bergamasco and Malanotte–Rizzoli, 2010; Tsimplis et al., 2006). These intermediate waters are formed by convection in the south of Rhodes (LIW) and in the Aegean Sea (Cretan intermediate water) and are found along the whole Mediterranean Sea between 200–500 m depth (Roether et al., 1998; Tsimplis et al., 2006). They present the maximum salinity of the Mediterranean water masses and outflow at the Strait of Gibraltar (Emelianov et al., 2006). The EMDW is formed in the Ionian Sea when water from the Southern Adriatic Sea plunges down through the Strait of Otranto and sinks to depths >3000 m. Then it flows eastwards and occupies the water column below the LIW in the eastern Mediterranean basin presenting potential temperatures >13.3°C and salinities >38.66 (Wu et al., 2000). For a short period of time, during the Eastern Mediterranean Transient (EMT) in the middle 1990's, the main deep water formation area was the Aegean Sea due to an abrupt shift in the climate and hydrography in this area, providing a warmer, more saline and denser deep water mass than the previously existing EMDW of Adriatic origin (Roether et al., 1996; Lascaratos et al., 1999; Klein et al., 2003). Hence, pre– and post–EMT varieties of EMDW of Adriatic origin coexists in the bathypelagic layer of the eastern basin. On the other basin, the WMDW is formed in winter in the Gulf of Lions (Gascard, 1978) and occupies the water column below the LIW in the western Mediterranean basin with temperatures between 12.75 and 12.80°C and salinities between 38.44 and 38.46 (Millot, 1999). During the winter of 2004–2005, a strong convection event in the Gulf of Lions (Western Mediterranean Transition) led to the formation of a new WMDW variety, saltier and slightly warmer than previously (salinity of 38.47–38.50 and temperature of 12.87–12.90ºC compared to 38.41–38.47 and 12.75–12.92ºC; López–Jurado et al., 2005; Beuvier et al., 2012; Schroeder et al.,

2016). Therefore, in the bathypelagic layer of the western Mediterranean, different varieties of WMDW coexist as well.

Sampling and determination of core parameters

Water samples were collected during the trans–Mediterranean cruise HOTMIX aboard R/V Sarmiento de Gamboa in spring 2014 (Heraklion, Crete, 27 April – Las Palmas, Canary Islands, 29 May). The transect consisted of 24 stations crossing the Mediterranean Sea from the Levantine Basin to the Strait of Gibraltar and 5 stations in the adjacent Northeast Atlantic Ocean (Figure 1). At each station, full–depth continuous conductivity–temperature–depth (SBE 911 plus CTD probe), dissolved oxygen (SBE–43 oxygen sensor) and chlorophyll fluorescence (SeaPoint fluorometer) profiles were recorded. Water samples were collected to analyse salinity, dissolved oxygen (DO) and chlorophyll *a* (Chl *a*) to calibrate the sensors for conductivity, DO and fluorescence, respectively. Conductivity measurements were converted into practical salinity scale values (UNESCO, 1985). Samples for salinity were measured with a Guildline Portasal salinometer Model 8410A. Chl *a* concentration was determined in seawater samples (500 mL) filtered through Whatman GF/F filters and stored frozen until analysis. Pigments were extracted in cold acetone (90% v/v) for 24 h and analyzed by means of a 10 AU Turner Designs bench fluorometer, previously calibrated with pure Chl *a* (Sigma Aldrich), according to Holm–Hansen et al. (1965). Dissolved oxygen was determined following the Winkler potentiometric method modified after Langdon (2010). The apparent oxygen utilization ($AOU = O_2sat - O_2$) was calculated using the algorithm proposed by Benson & Krause (UNESCO, 1986) for oxygen saturation (O_2sat). Potential temperature (θ) was calculated using TEOS–10 (UNESCO, 2010)

166 **Collection of SPE–DOM samples**

167 At nine stations (red asterisks in Figure 1) water samples were collected for the
168 solid phase extraction of DOM (SPE–DOM) to perform FT–ICR–MS analysis. Four to
169 five depths were sampled depending on the bathymetry of the stations (Figure 2), except
170 for the site at the Strait of Sicily (stn 13) where only LIW was sampled due to its
171 shallowness. The deep chlorophyll maximum (DCM) was sampled according to the
172 maximum fluorescence intensity, the LIW was sampled at the absolute maximum of the
173 salinity profile at each station, the oxygen minimum layer (OML) was established on
174 basis of the absolute minimum of the dissolved oxygen profile, and the deep waters
175 were sampled according to the salinity and temperature characteristic of the
176 bathypelagic zone of the eastern and western Mediterranean basins (Figure 2).

177 The sampling strategy was restricted by time constraints during the cruise, which
178 limited the number of collected samples, especially in the bathypelagic layer where
179 different varieties of deep waters were found. Unfortunately, we were not able to
180 sample any station in the Ionian Sea, so we missed the youngest variety of the EMDW.
181 Note that although the hydrographic properties of the OML sample at stn 4 are reported
182 in Table S1, this sample was rejected from the FT–ICR–MS analysis due to a
183 contamination problem.

184 Water samples were collected in 5–litres acid–cleaned polycarbonate carboys, and
185 then stored in the dark at 13°C until filtration within 5 hours. Filtration was performed
186 through precombusted (450°C, 4 h) Whatman GF/F filters in an acid–clean all–glass
187 filtration system under positive pressure with low flow of high purity N_2. Two–litres
188 aliquots of the filtrate were collected in acid–cleaned PTFE bottles for SPE–DOM.
189 Approximately 10 mL of the filtrate were collected for initial DOC determination in
190 precombusted (450°C, 12 h) glass ampoules. These samples were acidified with H_3PO_4

(85%, p.a., Merck) to pH < 2, the ampoules were fire–sealed and stored in the dark at 4°C until analysis in the base laboratory. DOC concentrations were measured with a Shimadzu TOC–V organic carbon analyser by high temperature catalytic oxidation (HTCO). The system was calibrated daily with potassium hydrogen phthalate (99.95–100.05%, p.a., Merck). The precision of the equipment was ±1 μmol L^{-1}. The accuracy was successfully tested daily with the DOC reference materials provided by D. A. Hansell (University of Miami, USA).

For SPE–DOM isolation, the filtered sea water sample (2 L) was acidified to pH 2 with HCl (37%, p.a., Merck) and the DOM was extracted on board with commercially available modified styrene divinyl benzene polymer cartridges (PPL, Agilent) as described in Dittmar et al. (2008). After extraction, cartridges were rinsed with acidified ultrapure water (pH 2, HCl 37%, p.a., Merck) to remove remaining salts and frozen at –20°C. Once in the base lab, the cartridges were dried by flushing with high purity N_2 and eluted with 6 mL of methanol (HPLC–grade, Sigma–Aldrich). Extracts were stored in amber vials at –20°C. DOC concentrations in the extracts were measured after complete evaporation of an aliquot and re–dissolution in ultrapure water. The extraction efficiency is the ratio of SPE–DOC to initial DOC concentrations. The mean extraction efficiency was 47.3 ± 3.9% on a carbon basis. Some of the epipelagic water samples showed slightly lower extraction efficiencies, likely due to the fact that PPL cartridges do not efficiently elute/retain the larger molecules (Raeke et al., 2016; Chen et al., 2016), which can be a significant fraction of DOM in the surface mixed layer

FT–ICR–MS analysis

SPE–DOM methanol extracts were diluted with ultrapure water and methanol (MS grade) to yield a DOC concentration of 15 mg C L^{-1} and a methanol–to–water ratio of 1:1 (v/v). Duplicates of each sample were prepared for analysis by ultrahigh–

resolution mass spectrometry using a Solarix FT–ICR–MS (Bruker Daltonik GmbH) connected to a 15 Tesla superconducting magnet (Bruker Biospin). Samples were infused at a flow rate of 120 μL h^{-1} into the electrospray source (ESI; Apollo II ion source, Bruker Daltonik GmbH) with the capillary voltage set to 4 kV in negative mode. Ions were accumulated in the hexapole for 0.3 s prior to transfer into the ICR cell. Data acquisition was done in broadband mode with a scanning range of 150–2000 Da. For each mass spectrum, 500 scans were accumulated. The spectra were mass calibrated (linear) using the Bruker Daltonics Data Analysis software package with an internal calibration list consisting of 51 known $C_xH_yO_z$ molecular formulae over the mass range of the samples. With this calibration procedure, a mass error of <0.1 ppm was achieved. SPE–DOM from the North Equatorial Pacific Intermediate Water (NEqPIW) collected at a depth of 670 m at the Natural Energy Laboratory of Hawaii Authority (NELHA) in Kona, Hawaii (Green et al., 2014) was used as an internal reference sample to assess instrument variability over time (Osterholz et al., 2014; Hansman et al., 2015). Molecular formulae were assigned to peaks considering a maximum mass error of 0.5 ppm and in the mass range between 150 and 850 Da by applying the following restrictions: $^{12}C_{1–130}{}^{1}H_{1–200}\,O_{1–50}{}^{14}N_{0–4}S_{0–2}P_{0–2}$ as described in Seidel et al. (2014). Only compounds with a signal–to–noise (S/N) ratio of 4 and higher were used for further analysis. Moreover, compounds present in less than 20% of samples with a maximum S/N less than 20 were removed, as well as the molecules containing the following heteroatom combinations: NSP, N_2S, N_3S, N_4S, N_2P, N_3P, N_4P, NS_2, N_2S_2, N_3S_2, N_4S_2 and S_2P as these are less likely to occur in nature and, furthermore, to be more conservative in assigning molecular formulae to a given *m/z*. The FT–ICR–MS signal intensity of each identified molecular formula was normalized to the sum of all molecular formula intensities with S/N higher than 5 in each sample. We assumed that

241 the inorganic (and organic) matrix is approximately the same for all the samples, so the
242 intensity of each molecular mass is only affected by its concentration. Further, Seidel et
243 al. (2015) incrementally mixed Amazon DOM with open Atlantic Ocean DOM showing
244 that the response signal of ESI–FT–ICR–MS was linear to the mixing ratio. Therefore,
245 we interpret the FT–ICR–MS data semi–quantitatively (Seidel et al., 2015; Hawkes et
246 al., 2016). The analytical window of the FT–ICR–MS was restricted by both the SPE
247 method and electrospray ionization efficiency. The former, using PPL cartridges, allows
248 to concentrate from the most apolar DOM species through to highly polar molecules,
249 but not the smallest polar molecules (i.e. short chain organic acids and free amino acids)
250 and colloidal aggregates (Hawkes et al., 2016; Raeke et al., 2016; Chen et al., 2016).
251 ESI is a low–fragmentation technique that preferentially ionizes polar functional groups
252 (Kujawinski, 2002), therefore carbohydrates are likely less efficiently ionized by ESI
253 than organic acids as it was suggested by Stubbins et al. (2010).

254 The aromaticity and the degree of unsaturation of a compound were assessed
255 based on its molecular formula and were expressed as the modified aromaticity index
256 ($AI_{mod} = (1 + C - {}^{1}/_{2}O - S - {}^{1}/_{2}H - {}^{1}/_{2}N - {}^{1}/_{2}P) / (C - {}^{1}/_{2}O - S - N - P)$) and double
257 bond equivalents ($DBE = 1 + {}^{1}/_{2}(2C - H + N + P)$), respectively (Koch and Dittmar,
258 2006, 2016). Higher aromaticity index and DBE are indicative of higher presence of
259 aromatic or even condensed aromatic molecules (Koch and Dittmar, 2006), which have
260 been suggested to be resistant to biodegradation (Stubbins et al., 2010; Rossel et al.,
261 2013). The degradation index (Ideg) was calculated using the formula proposed by
262 Flerus et al. (2012) ranging between 0 and 1. It is used as a simple proxy to assess the
263 relative degradation state of the SPE–DOM, and Flerus et al. (2012) suggested that a
264 higher Ideg points towards a more reworked DOM. The intensity–weighted averages of
265 molecular weight, number of elemental atoms (C, H, O), number of heteroatoms (N, S,

P), molar ratios (H/C, O/C and C/N), AI_{mod} and DBE were calculated for each sample by taking into account the FT–ICR–MS signal intensity of each assigned molecular formula. We sorted the assigned formulae into groups of formulae containing the following atoms: CHO, CHON, CHOS, CHOP, CHONS and CHOSP. In addition, we assigned the identified molecular formula to compound groups based on established molar ratios, AI_{mod}, DBE and heteroatoms contents (Seidel et al., 2014). The compound groups used in this work were: i) polyphenols ($0.5 < AI_{mod} < 0.666$) which are highly aromatic compounds, ii) highly unsaturated compounds ($AI_{mod} < 0.5$, $H/C < 1.5$ and $O/C < 0.9$), iii) unsaturated aliphatic ($1.5 < H/C < 2$, $O/C < 0.9$ and $N = 0$) and iv) carboxyl–rich alicyclic molecules (CRAM, $0.3 < DBE/C < 0.68$, $0.2 < DBE/H < 0.95$ and $0.77 < DBE/O < 1.75$) as described by Hertkorn et al. (2006). These parameters are summarized in Table S4. Note that 94% of the molecular formulae assigned to CRAM were also classified as highly unsaturated compounds. As this grouping includes a mixture of structural isomers and does not imply the presence of a structural entity in the sample (Seidel et al., 2014), we emphasize that this categorization is not unambiguous and alternative structures may exist for a given molecular formula. However, this classification is a useful tool to identify likely structures behind an identified molecular formula. All molecular parameters of each sample were calculated as averages of the duplicates.

Statistical analysis

Bray–Curtis dissimilarity matrices (Bray and Curtis, 1957) were computed based on relative signal average intensities. Principal Coordinates Analysis (PCoA) was then used for graphical representation of the DOM variability on the first two major axes of compositional change. Environmental and calculated parameters were correlated to the PCoA factors and graphed accordingly (Pearson's product moment correlation). The

analyses were performed in R (version 3.1.1, R Development Core Team 2012, [http://cran.r–project.org/]) and using the package vegan (Oksanen et al., 2016).

Multiple linear regressions were performed using R. Moreover, the Student's t–test was used for determining the significant differences between sample means (Table S3).

Results

Hydrography and bulk dissolved organic carbon background

Discrete sampling depths were chosen on basis of the vertical profiles of S, θ, DO and Chl *a* (Figure 2). The DCM (black dots in Figure 2) was deeper in the eastern than in the western basin (Table S1), showing higher DO and Chl *a* in the western basin. Regarding the LIW, the salinity maximum (yellow dots in Figure 2) was found between 200–300 m in the eastern basin, accompanied by a relative maximum of θ and DO. In the western basin the salinity maximum was located deeper (between 350–400 m) concurring with a relative maximum of θ and a minimum of DO. In general, the S, θ and DO along the core of the LIW were lower in the western than in the eastern basin (Table S1, Figure 2a–c). The oxygen minimum layer (green dots in Figure 2a–c) was found at 744 ± 211 m (n = 4) in the eastern basin. It coincided with the depth of the LIW in the western basin. Regarding the bathypelagic layer (red dots in Figure 2), the eastern basin was dominated by the EMDW, which was saltier and warmer than the analogous waters in the western basin, dominated by the WMDW. DO values were similar in both basins.

DOC concentrations (Figure S1 and Table S1) showed the maximum values of the sampling depths at the DCM (>60 μmol L^{-1}), decreasing to a minimum of 43–44 μmol L^{-1} in the deep waters. In the epipelagic layer an inverse relationship between Chl *a* and DOC concentration was observed (Figure 1d and S1a). It is remarkable that in the LIW

316 the DOC decreased significantly from 60.2 ± 0.9 μmol L^{-1} in the easternmost station
317 (stn 1) to 47.4 ± 0.7 μmol L^{-1} in the western basin (stn 18) (Figure S1). Conversely, in
318 the deep waters the distribution of the DOC did not reveal any significant gradient
319 between the western and eastern basins. Our DOC concentrations confirmed the
320 published vertical profile in both basins (Pujo–Pay et al., 2011; Santinelli, 2015 and
321 references therein). Atlantic samples showed similar DOC concentrations as the
322 Mediterranean Sea samples (Figure S1).

323 **Mediterranean SPE–DOM molecular signatures**

324 The distribution of the concentration of solid phase extracted DOC (SPE–DOC)
325 was parallel to the concentration of the bulk DOC (Figure S1) although carbon
326 extraction efficiency was 47.3 ± 3.9%. SPE–DOC therefore constitutes a good proxy of
327 the bulk DOC.

328 A total of 6057 resolved molecular masses of singly charged compounds were
329 detected in the FT–ICR–MS spectra of the 32 SPE–DOM samples analysed from the
330 Mediterranean Sea and Northeast Atlantic Ocean, covering a mass range of 154–817
331 Da. We identified 3689 molecular formulae in the mass range of 157–736 Da, not
332 considering ^{13}C isotopologues. The most abundant type of formulae was CHO, followed
333 by CHON, CHOS, CHOP, CHONS and finally CHOSP (Table S3).

334 The NEqPIW sample repeatedly analysed as a reference sample to control the
335 instrument variation over time also let us compare the molecular composition of the
336 SPE–DOM in the Mediterranean Sea with one of the oldest water masses of the world
337 ocean: NEqPIW (Table S3). As expected, we observe that the Mediterranean Sea
338 contains significantly less reworked DOM (lower molecular weight, O/C, DBE and
339 Ideg) than the NEqPIW.

The Atlantic water entering the Mediterranean Sea through the Strait of Gibraltar (represented by the DCM sample at stn 25; black dot in Figure 3) exhibited a significantly different molecular composition than the overflow of Mediterranean water (represented by the LIW at stn 18; yellow dot in Figure 3). Specifically, the SPE–DOM found in the Atlantic water inflow displayed lower molecular weight, O/C ratio, DBE, Ideg, CRAM and highly unsaturated compounds contribution, as well as a higher proportion of unsaturated aliphatic compounds and an increased H/C ratio (Figure 3 and Table S2). A principal coordinates analysis (PCoA) to link molecular composition to environmental parameters that includes all samples (Figure 4) also reveals the molecular dissimilarity between the Atlantic inflow and the Mediterranean overflow. While the Atlantic sample was found in the negative part of both axes, the LIW sample was found in the positive. Note that the first 2 coordinates of the PCoA comprised 72% of the SPE–DOM molecular variability. The differences between the inflow and outflow at the Strait of Gibraltar were also explored using a differential mass spectrum (Figure 5a), subtracting the normalized peak intensities of the LIW at stn 18 from the normalized peak intensities of the DCM at stn 25. Positive differences of intensity showed peaks of higher relative intensities in the Atlantic inflow with an average molecular weight of 349 Da. Negative signals indicated peaks of higher relative intensities in the Mediterranean overflow enriched in compounds of an average molecular weight of 432 Da. Note that LIW at stn 18 was chosen to represent the Mediterranean water overflow as LIW at stn 21 is influenced by the mixing with WMDW as it is located close to the deep waters in the θ/S diagram (Figure S2c) and apart from the LIW at stns 15 and 18 in the PCoA (Figure 4).

To study the DOM degradation along the shallow overturning circulation cell of the Mediterranean Sea we followed the compositional changes of SPE–DOM collected

365 in the core–of–flow of the LIW. The molecular characteristics of the LIW samples were
366 uneven and the samples were split in the PCoA (Figure 4): while the eastern basin
367 samples were found in the negative part of the first coordinate, the western basin
368 samples were located in the positive. In addition, a constrained analysis of principal
369 coordinates (CAP) based on Bray–Curtis dissimilarities revealed significant molecular
370 differences between both basins ($p < 0.05$, n=8). Specifically, as the LIW flowed
371 westwards we observed a raise of molecular size (Figure 3a), O/C ratio, DBE and
372 AImod. Moreover, we observed an increase of the Ideg and the proportion of highly
373 unsaturated molecules, as well as a decrease of the proportion of unsaturated aliphatic
374 molecules (Figure 3b–d and Table S2). Again, these differences were examined in more
375 detail using a differential spectrum (Figure 5b), subtracting the normalized peak
376 intensities of the LIW at stn 18 (western basin, more reworked DOM) from the
377 normalized peak intensities of the LIW at stn 1 (eastern basin, fresher DOM). Positive
378 differences of intensity showed peaks with higher relative intensities in the LIW at stn
379 1, where this sample presented an enrichment of molecules with an average molecular
380 weight of 335 Da. Negative signals indicated peaks with higher relative intensities in
381 the LIW at stn 18, sample enriched in molecules of an average molecular weight of 427
382 Da.

383 Concerning the deep waters, in the eastern basin samples collected at 2000 m
384 depth at stns 1 and 4 (EMT) fell closely together in the PCoA analysis (Figure 4).
385 However, the samples collected at 1000 m at stns 1 and 4 (pre–EMT) and at 2000 m at
386 stn 7 (post–EMT) grouped in our ordination and were separated from the EMT samples.
387 Comparing the molecular composition of SPE–DOM in these water masses, the deep
388 water at stn 7 showed a less degraded SPE–DOM signature (lower molecular weight,

389 oxygen, DBE, Ideg, highly unsaturated and higher contribution of unsaturated aliphatic
390 compounds; Figure 3 and Table S2).

391 Globally, the PCoA revealed a clear separation of the samples by water layers.
392 Along both PCoA axes, less reworked DOM was clearly separated from more degraded
393 DOM. Samples collected at the DCM (recently formed) were characterized by less
394 reworked DOM (higher proportion of unsaturated aliphatic compounds, higher H/C
395 ratio and DOC concentration). Highly unsaturated molecules, likely susceptible to
396 photochemical processes, and photoresistant unsaturated aliphatic molecules comprised
397 the largest fraction of the SPE–DOM (>90%) in all samples (Table S2). Since
398 abundances are expressed as relative contributions (percentages), the increase of one
399 kind of compounds entails the decrease of the others. Thus, the lower percentage of
400 highly unsaturated compounds in the epipelagic layer (Figure 3b) could be due to the
401 photochemical removal of this type of molecules in the photic layer.

402 **Discussion**

403 Despite the relatively small size and short residence times of the Mediterranean
404 Sea, a PCoA analysis revealed that the molecular composition of DOM was found to be
405 heterogeneous among basins and water layers. We show that this heterogeneity is
406 mainly caused by three different factors that control the DOM molecular composition in
407 the Mediterranean Sea: water mass origin, biodegradation and photobleaching.

408 **Water mass origin as a driver of SPE–DOM composition**

409 We found that the overall molecular composition of the Mediterranean Sea was
410 significantly less reworked than the NEqPIW (Table S3). This water mass acts a
411 reference for refractory DOM, as it is one of the oldest, least ventilated, water masses of
412 the global ocean (Stuiver et al., 1983; Osterholz et al., 2015). This result demonstrates

413 the noticeable molecular differences between the Mediterranean Sea DOM and the
414 NEqPIW DOM in spite of presenting similar radiocarbon ages as recently reported by
415 Santinelli et al. (2015).

416 The exchange across the Strait of Gibraltar results in the Atlantic inflow
417 transporting relatively fresh, nutrient–poor (Huertas et al., 2012), and less degraded
418 DOC–rich surface Atlantic waters into the Mediterranean Sea. On the other hand, the
419 Mediterranean outflow transports salty, nutrient–rich and reworked DOC–poor
420 intermediate Mediterranean water into the Atlantic. Accordingly, Ideg and DBE of the
421 SPE–DOC in the Mediterranean overflow increased by 25% and 5%, respectively and
422 the proportion of unsaturated aliphatic compounds decreased by 24% compared to the
423 Atlantic inflow. Given that the Atlantic water that enters the Mediterranean Sea is part
424 of the shallow overturning cell of the Mediterranean intermediate waters that constitute
425 the Gibraltar overflow (Schneider et al., 2014), it is expected that the difference in
426 composition is implemented during completion of that overturning circulation within
427 the Mediterranean Sea basin. When the LIW is formed (stn 1) the DOM molecular
428 composition is similar to the DOM Atlantic inflow (DCM in stn 25) (stn 25– black dot
429 and stn 1–yellow dot in Figure 4). However, once the LIW flows in the shallow
430 overturning circulation across the Mediterranean Sea this water mass transports more
431 degraded DOM as will be discussed in the next section.

432 Deep waters in the eastern basin exhibited significant molecular differences
433 related to their formation site (Aegean *versus* Adriatic) or time (pre– *versus* post–EMT).
434 Samples collected at 2000 m depth at stns 1 and 4 presented more degraded DOM
435 (molecular weight increased by 5 Da, the proportion of polyphenols by 6% and the
436 proportion of unsaturated aliphatic compounds decreased by 9%) than samples collected
437 at 1000 m at the same stations (Figure 3, 4, Table S2). These molecular changes were

attributed to the fact that samples collected at 2000 m were formed in the Aegean Sea during the EMT (according to their thermohaline properties and oxygen concentrations; Figure S2d and Table S1; Roether et al., 1996, Lascaratos et al., 1999, Klein et al., 2003). However, samples collected at 1000 m were formed in the Adriatic Sea in the pre–EMT which showed lower S and DO concentration than the deep waters of Aegean origin (Lascaratos et al., 1999; Schneider et al., 2014). These results support the idea that water mass origin drives the DOM molecular composition. Regarding the formation time of these deep waters it would be plausible that the oldest (pre–EMT) would host most degraded DOM, however the results show the opposite. It seems that water mass origin has a higher impact on DOM molecular composition than water mass age in the deep waters of the eastern Mediterranean Sea. On the other hand, samples collected at 2000 m at stn 7 (post–EMT) presented less degraded DOM (lower molecular weight (–5 Da), Ideg (–7%), DBE (–2%) and an about 9% higher proportion of unsaturated aliphatic compounds) than the sample collected at stn 1 at 1000 m (pre–EMT). In this case, although both water masses were formed in the Adriatic Sea, their formation times were different according to their thermohaline properties and DO concentrations (Figure S2d and Table S1), which lead to dissimilar DOM molecular signatures.

Diagenetic transformations of SPE–DOM through the Mediterranean Sea

The samples taken along the salinity maximum of the LIW show a total decrease in DOC concentration of 15 μmol L^{-1} attributed to prokaryotic degradation within the shallow overturning circulation of the Mediterranean Sea (black dashed line in Figure 2a–c and 3). A highly significant linear relationship between DOC and AOU was found along the LIW pathway followed during the cruise ($r = –0.96$, $p < 0.001$, $n = 9$). The slope of this linear regression (model II; Sokal and Rohlf, 1995) was –0.21 ± 0.02, which, converted into oxygen equivalents using the canonical Redfield $–O_2/C$ ratio of

463 1.4, translates into 0.30 ± 0.03, i.e. the 30 ± 3% of the oxygen consumption was due to
464 the microbial oxidation of DOC. To minimize the effect of water mass mixing on the
465 DOC/AOU relationship, we performed a multiple linear regression of DOC with θ, S
466 and AOU. Doing this the slope changed to –0.31 ± 0.08, which, converted into oxygen
467 equivalents as above, resulted in a 43 ± 11% of the oxygen consumption due to
468 prokaryotic oxidation of DOC. Note that this number is not significantly different from
469 the previously obtained with the simple linear regression. These values are consistent
470 with previous estimates in the Mediterranean Sea by Santinelli et al. (2010, 2012),
471 which ranged from 38% to 53%. Such a result is much higher than the 10–20% found in
472 the dark global ocean (Arístegui et al., 2002). Warmer deep water temperatures in the
473 Mediterranean Sea (>13°C *versus* <5°C; Dickson and Brown, 1994), which stimulate
474 the prokaryotic degradation processes, is the likely reason behind this difference.
475 Microbial oxidation leads to decreasing DOC concentrations along the overturning cell
476 and the remaining DOC is, indeed, more reworked. In this regard, in agreement with
477 Hansman et al. (2015), we found a positive relationship between the Ideg and AOU for
478 all collected samples: Ideg = (0.0019 ± 0.0002) AOU + (0.50 ± 0.01) (R^2 = 0.71, p <
479 0.0001). The obtained regression slope suggests that the degradation ratio per oxygen
480 consumption unit is faster in the Mediterranean Sea than in the Atlantic Ocean, given
481 that it is significantly higher (p < 0.0001) than the regression slope obtained by
482 Hansman et al. (2015) in the Atlantic Ocean (slope = 0.00073 ± 0.00003, R^2 = 0.64, p <
483 0.001). This more efficient degradation could also be related to the above mentioned
484 warmer deep water temperatures in the Mediterranean Sea. As a consequence, the rate
485 of microbial processes would be about 2–fold higher in the Mediterranean Sea than in
486 the Atlantic Ocean, according to the Arrhenius law. As for the case of DOC, to assess
487 the role of water mass mixing in this relationship we performed a multiple linear

488 regression of Ideg with θ, S and AOU (R^2 = 0.80; p < 0.003, Ideg/AOU slope = 0.0011
489 ± 0.0003). Whereas the standard deviation of Ideg (SD_{Ideg} = 0.065) retains the
490 variability due to both water mass mixing and biogeochemical processes in the
491 Mediterranean Sea, the standard deviation of the residuals of the multiple linear
492 regression of Ideg with θ and S ($SD_{\Delta Ideg}$ = 0.035) retains only the variability due to
493 biogeochemical processes. Therefore, from the ratio of both SD, it can be inferred that
494 54% of the observed variability of Ideg could be explained by processes not associated
495 to water mass mixing. We could presume that the mixing of LIW, with EMDW in the
496 eastern basin and with WMDW in the western basin, can be partly responsible for the
497 different molecular composition observed in the LIW of stns 1 and 18. However, in the
498 western basin, mixing cannot be the only process affecting the LIW, since this water
499 mass with the highest AOU is surrounded by water bodies (the DCM on top and the
500 WMDW underneath) which present lower AOU values (Table S1). Concomitantly, part
501 of the observed changes in the molecular composition of SPE–DOM could be due to the
502 production/consumption of higher/lower molecular weight compounds as the LIW
503 becomes older (more degraded) in its route westwards. A parallel increase in molecular
504 weight (Figure 5b) and oxygenation (5%) and a decrease of the H/C ratio (Table S2)
505 were found to be indicators of degraded organic matter (Hertkorn et al., 2013, Chen et
506 al., 2014 and Flerus et al., 2012), in accordance with an increase of the Ideg (22%). A
507 higher DBE and AI_{mod} are indicative of an increasing degree of aromaticity and
508 unsaturation (Koch and Dittmar, 2006, 2016). In addition, a higher proportion of highly
509 unsaturated compounds (7%) is also indicative of reworked DOM, as these compounds
510 are considered refractory and produced during the remineralisation processes in the
511 meso– and bathypelagic layers (Seidel et al., 2015). Conversely, unsaturated aliphatic
512 compounds are considered bio–labile molecules, as they comprise a major fraction of

513 phytoplankton exudates (Medeiros et al., 2015). We observed a decrease by 34% of the
514 proportion of unsaturated aliphatic compounds in the LIW along its route westwards.

515 **Photodegradation versus biodegradation in the epipelagic layer**

516 Photochemical processes have been proposed as an abiotic pathway for DOC
517 degradation in the surface ocean (Mopper et al., 2015). The Ideg is used to assess the
518 degradation state of the SPE–DOM. Since photodegradation can produce bio–labile
519 aliphatic and peptide–like compounds (Stubbins and Dittmar 2015; Stubbins et al.,
520 2010), we hypothesize that this process could lead to a lower Ideg. We observed an
521 increasing trend of Ideg with depth (Figure 3d), likely indicating a synergy between an
522 increasing contribution of prokaryote DOM degradation with depth, and the potential
523 photodegradation and new production in the photic layer. However both processes, new
524 production and photodegradation, cannot be deciphered by applying the Ideg.
525 Therefore, we do not find conclusive evidence for the effect of photodegradation in the
526 Mediterranean Sea, probably due to the great depth of the shallowest level that we
527 sampled (DCM). In addition, the cruise was conducted in April–May, i.e. after winter
528 mixing and when solar radiation is not at the summer maximum. Further
529 studies/experiments should be performed to clarify the role of photochemical processing
530 on the DOM composition in the Mediterranean Sea.

531 **Conclusions**

532 Despite the small size and relatively short residence time of the Mediterranean
533 Sea, water mass origin and mineralization processes lead to contrasting molecular
534 composition of SPE–DOM with depth and basin. SPE–DOM in the Mediterranean Sea
535 was remarkably different from the SPE–DOM in the Atlantic Ocean inflow.
536 Considering the shallow overturning cell of the Mediterranean Sea, the evolution of the

molecular composition of SPE–DOM from the Levantine basin to the Strait of Gibraltar evidences the transformation of these materials since LIW is formed. As a result, a westward decrease of DOC concentrations and a lower proportion of unsaturated aliphatic compounds are observed, as well as an increase in average molecular weight and enrichment in unsaturation, oxygenation, state of degradation and highly unsaturated compounds as the SPE–DOM is degraded. We found that the water mass origin and the formation time lead to distinct DOM molecular properties. Thus, pre–EMT deep waters formed in the Adriatic Sea presented less degraded DOM than deep waters formed in the Aegean Sea during the EMT. In addition, different varieties of deep waters formed in the Adriatic Sea (pre– and post–EMT) presented different DOM molecular composition in spite of being formed in the same area. Consequently, the Mediterranean Sea constitutes a suitable model basin for future DOM studies as water bodies of different molecular composition can be observed in closest proximity. Taking advantages of forthcoming hydrographic cruises, it would be worthwhile to study the molecular composition of DOM in the Mediterranean waters at their formation sites and help to complete the picture of DOM molecular composition and turnover in the Mediterranean Sea.

Acknowledgements

The authors are grateful to the Captain, crew, technicians and scientists aboard the R/V Sarmiento de Gamboa for their support during the cruise. We especially thank M.J. Pazó, V. Vieitez, M. Friebe and I. Ulber for DOC measurements, K. Klaproth for support with FT–ICR–MS analysis, M. Manecki and B.E. Noriega for their help with data processing and J. Niggemann for valuable discussions. This work was financed by the project HOTMIX (grant number CTM2011–30010–C02–01–MAR and 02–MAR) and the project FERMIO (MINECO, CTM2014-57334-JIN), both co–financed with

562 FEDER funds. A.M.M.–P. was funded by a predoctoral fellowship (reference BES–
563 2012–056175) and a short stay fellowship (reference EEBB–I–14–08926) from the
564 Spanish Ministry of Economy and Competitiveness. M.N.–C. was partially supported
565 by the CSIC Program “Junta para la Ampliación de Estudios”, co–financed by the ESF
566 (reference JAE DOC 040), co–financed with FEDER funds.

References

Amon, R. M. and R. Benner. 1996. Bacterial utilization of different size classes of dissolved organic matter. Limnol. Oceanogr. **41:** 41–51.

Arístegui, J., C. M. Duarte, S. Agustí, M. Doval, X. A. Álvarez–Salgado, and D. A. Hansell. 2002b. Oceanography: Dissolved organic carbon support of respiration in the dark ocean. Science **298**: 1967

Benner, R., B. Biddanda, B. Black and M. McCarthy. 1997. Abundance, size distribution, and stable carbon and nitrogen isotopic compositions of marine organic matter isolated by tangential–flow ultrafiltration. Mar. Chem. **57:** 243–263.

Benner, R., J. D. Pakulski, M. McCarthy, J. I. Hedges and P. G. Hatcher. 1992. Bulk chemical characteristics of dissolved organic matter in the ocean. Science. **255:** 1561–1564.

Bergamasco, A. and P. Malanotte–Rizzoli. 2010. The circulation of the Mediterranean Sea: a historical review of experimental investigations. Adv. Oceanogr. Limnol. **1:** 11–28.

Béthoux, J.P., P. Morin, C. Chaumery, O. Connan, B. Gentili, and D. Ruiz–Pino. 1998. Nutrients in the Mediterranean Sea, mass balance and statistical analysis of concentrations with respect to environmental change. Mar. Chem. **63:** 155–169.

Beuvier, J., K. Béranger, C. Lebeaupin Brossier, and others. 2012. Spreading of the Western Mediterranean Deep Water after winter 2005: Time scales and deep cyclone transport. **117.**

Bray, J. R. and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin, Ecological Monographs, Ecological Society of America. **27:** 326 – 349.

Carlson, C.A., 2002. Chapter 4 – Production and Removal Processes, p. 91–151. *In* Dennis A. Hansell and Craig A. Carlson [eds.], Biogeochemistry of Marine Dissolved Organic Matter. Academic Press, San Diego.

Chen, H., A. Stubbins, E. M. Perdue, N. W. Green, J. R. Helms, K. Mopper and P. G. Hatcher. 2014. Ultrahigh resolution mass spectrometric differentiation of dissolved organic matter isolated by coupled reverse osmosis–electrodialysis from various major oceanic water masses. Mar. Chem. **164:** 48–59.

Chen, M., S. Kim, J.–E. Park, H.–J. Jung and J. Hur. 2016. Structural and compositional changes of dissolved organic matter upon solid–phase extraction tracked by multiple analytical tools. Anal. Bioanal. Chem. **408:** 6249–6258.

Cruzado, A., 1985. Chemistry of Mediterranean waters, p. 126–147. *In* R. Margalef, [eds.], The Western Mediterranean. Pergamon Press, Oxford.

603 Dickson, R. R. and J. Brown. 1994. The production of North Atlantic Deep Water:
604 sources, rates, and pathways. J. Geophys. Res. **99:** 12 319–12 341.

605 Dittmar, T., B. Koch, N. Hertkorn and G. Kattner. 2008. A simple and efficient method
606 for the solid–phase extraction of dissolved organic matter (SPE–DOM) from seawater.
607 **6:** 230–235.

608 Emelianov, M., J. Font, A. Turiel, C. Mullot, J. Solé, P. M. Poulain, A. Julià and M. R.
609 Vitrià. 2006. Transformation of levantine intermediate water tracked by MEDARGO
610 floats in the Western Mediterranean. Ocean Sci. **2:** 281–290.

611 Flerus, R., O. J. Lechtenfeld, B. P. Koch, S. L. McCallister, P. Schmitt–Kopplin, R.
612 Benner, K. Kaiser, and G. Kattner. 2012. A molecular perspective on the ageing of
613 marine dissolved organic matter. **9:** 1935–1955.

614 Gascard, J.C., 1978. Mediterranean deep water formation; baroclinic instability and
615 oceanic eddies. Oceanol. Acta **1:** 315–330.

616 Green, N. W., E. M. Perdue, G. R. Aiken, K. D. Butler, H. Chen, T. Dittmar, J.
617 Niggemann and A. Stubbins. 2014. An intercomparison of three methods for the large–
618 scale isolation of oceanic dissolved organic matter. Mar. Chem. **161:** 14–19.

619 Hansell, D.A., 2002. Chapter 15 – DOC in the Global Ocean Carbon Cycle, p. 685–715.
620 *In* Dennis A. Hansell and Craig A. Carlson [eds.], Biogeochemistry of Marine
621 Dissolved Organic Matter. Academic Press, San Diego.

622 Hansell, D. A., C. A. Carlson, D. J. Repeta and R. Schlitzer. 2009. Dissolved organic
623 matter in the ocean a controversy stimulates new insights. Oceanography **22:** 202–211.

624 Hansman, R. L., T. Dittmar and G. J. Herndl. 2015. Conservation of dissolved organic
625 matter molecular composition during mixing of the deep water masses of the northeast
626 Atlantic Ocean. Mar. Chem. **177:** 288–297.

627 Hawkes, J. A., C. T. Hansen, T. Goldhammer, W. Bach, and T. Dittmar. 2016.
628 Molecular alteration of marine dissolved organic matter under experimental
629 hydrothermal conditions. Geochim. Cosmochim. Acta **175**: 68–85.

630 Hedges, J.I., 1992. Global biogeochemical cycles: progress and problems. Mar. Chem.
631 **39:** 67–93.

632 Hertkorn, N., R. Benner, M. Frommberger, P. Schmitt–Kopplin, M. Witt, K. Kaiser, A.
633 Kettrup and J. I. Hedges. 2006. Characterization of a major refractory component of
634 marine dissolved organic matter. Geochim. Cosmochim. Acta. **70:** 2990–3010.

635 Hertkorn, N., M. Harir, B. P. Koch, B. Michalke and P. Schmitt–Kopplin. 2013. High–
636 field NMR spectroscopy and FTICR mass spectrometry: Powerful discovery tools for
637 the molecular level characterization of marine dissolved organic matter.
638 Biogeosciences.**10:** 1583–1624.

639 Holm–Hansen, O., C. J. Lorenzen, R. W. Holmes and J. D. Strickland. 1965.
640 Fluorometric determination of chlorophyll. J. Cons. Int. Explor. Mer **30:** 3–15.

641 Huertas, I. E., A. F. Ríos, J. García–Lafuente, and others. 2012. Atlantic forcing of the
642 Mediterranean oligotrophy. Glob. Biogeochem. Cycles **26**.

643 Jones, V., T. B. Meador, A. Gogou, C. Migon, K. E. H. Penkman, M. J. Collins and D.
644 J. Repeta. 2013. Characterisation and dynamics of dissolved organic matter in the
645 Northwestern Mediterranean Sea. Prog. Oceanogr. **119:** 78–89.

646 Klein, B., W. Roether, N. Kress, B. B. Manca, and others. 2003. Accelerated oxygen
647 consumption in eastern Mediterranean deep waters following the recent changes in
648 thermohaline circulation. J. Geophys. Res. **108**.

649 Koch, B.P. and T. Dittmar. 2006. From mass to structure: An aromaticity index for
650 high–resolution mass data of natural organic matter. Rapid Commun. Mass Spectrom.
651 **20**: 926–932.

652 Koch, B.P. and T. Dittmar. 2016. Erratum: From mass to structure: An aromaticity
653 index for high–resolution mass data of natural organic matter. Rapid Commun. Mass
654 Spectrom. **30:** 250.

655 Koch, B. P., M. Witt, R. Engbrodt, T. Dittmar and G. Kattner. 2005. Molecular
656 formulae of marine and terrigenous dissolved organic matter detected by electrospray
657 ionization Fourier transform ion cyclotron resonance mass spectrometry. Geochim.
658 Cosmochim. Acta. **69:** 3299–3308.

659 Koprivnjak, J. F., P. H. Pfromm, E. Ingall, and others. 2009. Chemical and
660 spectroscopic characterization of marine dissolved organic matter isolated using
661 coupled reverse osmosis–electrodialysis. Geochim. Cosmochim. Acta. **73:** 4215–4231.

662 Kujawinski, E. B. 2002. Electrospray ionization Fourier transform ion cyclotron
663 resonance mass spectrometry (ESI FT–ICR MS): Characterization of complex
664 environmental mixtures. Environ. Forensics **3**: 207–216.

665 Langdon, C. 2010. Determination of dissolved oxygen in seawater by Winkler titration
666 using the amperometric technique. GO–SHIP repeat hydrography manual: a collection
667 of expert reports and guidelines edited by BM Sloyan and C.Sabine, IOC/IOCCP, Paris.

668 Laruelle, G. G., V. Roubeix, A. Sferratore, and others. 2009. Anthropogenic
669 perturbations of the silicon cycle at the global scale: Key role of the land–ocean
670 transition. Glob. Biogeochem. Cycles **23**.

671 Lascaratos, A., W. Roether, K. Nittis and B. Klein. 1999. Recent changes in deep water
672 formation and spreading in the eastern Mediterranean Sea: a review. Prog. Oceanogr.
673 **44:** 5–36.

674 Lechtenfeld, O. J., G. Kattner, R. Flerus, S. L. McCallister, P. Schmitt–Kopplin and B.
675 P. Koch. 2014. Molecular transformation and degradation of refractory dissolved

organic matter in the Atlantic and Southern Ocean. Geochim. Cosmochim. Acta **126:** 321–337.

López‐Jurado, J., C. Gonzalez‐Pola and P. Velez‐Belchi. 2005. Observation of an abrupt disruption of the long-term warming trend at the Balearic Sea, western Mediterranean Sea, in summer 2005. Geophys. Res. Lett. **32**.

Meador, T. B., A. Gogou, G. Spyres, and others. 2010. Biogeochemical relationships between ultrafiltered dissolved organic matter and picoplankton activity in the Eastern Mediterranean Sea. Deep–Sea Res. II Top. Stud. Oceanogr. **57:** 1460–1477.

Medeiros, P. M., M. Seidel, L. C. Powers, T. Dittmar, D. A. Hansell and W. L. Miller. 2015. Dissolved organic matter composition and photochemical transformations in the northern North Pacific Ocean. Geophys. Res. Lett. **42:** 863–870.

Millot, C. 1999. Circulation in the Western Mediterranean Sea. J. Marine Syst. **20:** 423–442.

Mopper, K. D., J. Kieber, and A. Stubbins. 2015. Marine photochemistry of organic matter: processes and impacts, p. 389–450. *In* Biogeochemistry of Marine Dissolved Organic Matter. Academic Press, Waltham MA, USA.

Oksanen, J., F. G. Blanchet, R. Kindt, and others. 2016. vegan: Community Ecology Package. R package version 2.3–5. http://cran.r–project.org.

Osterholz, H., J. Niggemann, H. Giebel, M. Simon and T. Dittmar. 2015. Inefficient microbial production of refractory dissolved organic matter in the ocean. Nat. Commun. **6:** 7422, 1–8.

Osterholz, H., T. Dittmar and J. Niggemann. 2014. Molecular evidence for rapid dissolved organic matter turnover in Arctic fjords. Mar. Chem. **160:** 1–10.

Pinardi, N. and E. Masetti. 2000. Variability of the large scale general circulation of the Mediterranean Sea from observations and modelling: A review. Palaeogeogr. Palaeoclimatol., Palaeoecol. **158:** 153–174.

Pujo–Pay, M., P. Conan, L. Oriol, and others. 2011. Integrated survey of elemental stoichiometry (C, N, P) from the western to eastern Mediterranean Sea. Biogeosciences **8:** 883–899.

Raeke, J., O. J. Lechtenfeld, M. Wagner, P. Herzsprung and T. Reemtsma. 2016. Selectivity of solid phase extraction of freshwater dissolved organic matter and its effect on ultrahigh resolution mass spectra. Environ. Sci.: Processes Impacts.

Roether, W., B. Klein, V. Beitzel and B. B. Manca. 1998. Property distributions and transient–tracer ages in Levantine Intermediate Water in the Eastern Mediterranean. J. Marine Syst. **18:** 71–87.

711 Roether, W., B. B. Manca, B. Klein, D. Bregant, D. Georgopoulos, V. Beitzel, V.
712 Kovacevic and A. Luchetta. 1996. Recent changes in eastern Mediterranean deep
713 waters. Science. **271:** 333–335.

714 Rossel, P. E., A. V. Vähätalo, M. Witt and T. Dittmar. 2013. Molecular composition of
715 dissolved organic matter from a wetland plant (Juncus effusus) after photochemical and
716 microbial decomposition (1.25 yr): Common features with deep sea dissolved organic
717 matter. Org. Geochem. **60:** 62–71.

718 Santinelli, C. 2015. Chapter 13 – DOC in the Sea, p. 579–608. *In* D. A. Hansell and C.
719 A. Carlson [eds.], Biogeochemistry of marine dissolved organic matter. Elsevier.

720 Santinelli, C., C. Follett, S. Retelletti Brogi, L. Xu, and D. Repeta. 2015. Carbon isotope
721 measurements reveal unexpected cycling of dissolved organic matter in the deep
722 Mediterranean Sea. Mar. Chem. **177**: 267–277. doi:10.1016/j.marchem.2015.06.018

723 Santinelli, C., L. Nannicini, and A. Seritti. 2010. DOC dynamics in the meso and
724 bathypelagic layers of the Mediterranean Sea. Deep–Sea Res. II: Top. Stud. Oceanogr.
725 **57:** 1446–1459.

726 Santinelli, C., R. Sempéré, F. Van Wambeke, B. Charriere and A. Seritti. 2012. Organic
727 carbon dynamics in the Mediterranean Sea: An integrated study. Glob. Biogeochem.
728 Cycles **26**.

729 Schlitzer, R. 2016. Ocean Data View. odv.awi.de.

730 Schneider, A., T. Tanhua, W. Roether, and R. Steinfeldt. 2014. Changes in ventilation
731 of the Mediterranean Sea during the past 25 year. Ocean Sci. 10: 1–16.

732 Schroeder, K., J. Chiggiato, H.L. Bryden, M. Borghini, S.B. Ismail. 2016. Abrupt
733 climate shift in the Western Mediterranean Sea. Sci. Rep. **6**.

734 Seidel, M., M. Beck, T. Riedel, and others. 2014. Biogeochemistry of dissolved organic
735 matter in an anoxic intertidal creek bank. Geochim. Cosmochim. Acta **140:** 418–434.

736 Seidel, M., P. L. Yager, N.D. Ward, and others. 2015. Molecular–level changes of
737 dissolved organic matter along the Amazon River–to–ocean continuum. Mar. Chem.
738 **177:** 218–231.

739 Sokal, F. F., F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in
740 biological research, 3rd ed. W. H. Freeman and Company, New York.

741 Stubbins, A., R. G. Spencer, H. Chen, and others. 2010. Illuminated darkness:
742 Molecular signatures of Congo River dissolved organic matter and its photochemical
743 alteration as revealed by ultrahigh precision mass spectrometry. Limnol. Oceanogr. **55:**
744 1467–1477.

745 Stubbins, A., T. Dittmar, 2015. Illuminating the deep: Molecular signatures of
746 photochemical alteration of dissolved organic matter from North Atlantic Deep Water**.**
747 Mar. Chem. **177:** 318–324.

748 Stuiver, M., P.D. Quay, H. G. Ostlund. 1983. Abyssal water carbon–14 distribution and
749 the age of the world oceans. Science. **219:** 849–851.

750 Tsimplis, M. N., V. Zervakis, S. A. Josey, and others. 2006. Chapter 4– Changes in the
751 oceanography of the Mediterranean Sea and their link to climate variability, p. 227–282.
752 *In* P. Lionello, P. Malanotte–Rizzoli, R. Boscolo, [eds.] Mediterranean Climate
753 Variability. Elsevier, Amsterdam.

754 UNESCO, 1985. The International System of Units (SI) in oceanography. UNESCO
755 Tech. Paper. Marine Science. **45:** 1–124.

756 UNESCO, 1986. Progress on oceanographic tables and standards 1983–1986. Work and
757 recommendations of UNESCO/SCOR/ICES/IAPSO joint panel UNESCO. Tech. Pap.
758 Marine Science. **50:** 1–59.

759 UNESCO, 2010. The international thermodynamic equation of seawater: Calculation
760 and use of thermodynamic properties. Intergovernmental Oceanographic Commission,
761 Manuals and Guides No.56, UNESCO/IOC/SCOR/IAPSO (English), pp 196.

762 Wu, P., K. Haines and N. Pinardi. 2000. Toward an understanding of deep–water
763 renewal in the Eastern Mediterranean. J. of Phys. Oceanogr. **30:** 443–458.

764

765 Figure 1: Study area and sampling stations. The circles depict all cruise stations and
766 asterisks represent the stations where samples were taken for molecular characterization
767 of the DOM. Figure created using ODV software (Schlitzer, 2016).

768

769 Figure 2: Distribution of (a) salinity (S), (b) potential temperature (θ) in ºC, (c)
770 dissolved oxygen (DO) in µmol/kg and (d) fluorescence of chlorophyll *a* (Chl *a*) in mg
771 m^{-3} obtained from the sensors attached to the rosette sampler along the Mediterranean
772 Sea. Black, yellow, green and red dots represent samples taken in the epipelagic layer
773 (DCM), Levantine Intermediate Water (LIW), oxygen minimum layer and deep waters,
774 respectively. The dashed black line represents the route of the LIW along the transect.
775 Note that the depth is displayed on a non-linear scale. Values from all stations were
776 used to show these distributions. Figure created using ODV software (Schlitzer, 2016).

777

778 Figure 3: Distributions of (a) molecular weight in Dalton, (b) highly unsaturated
779 compounds in %, (c) unsaturated aliphatic compounds in % and (d) degradation index
780 (range between 0 and 1, unitless) using averages of duplicates in the Mediterranean Sea
781 and Northeast Atlantic Ocean. The dashed black line represents the route of the LIW
782 along the transect. Note that the depth is displayed on a non-linear scale. Figure created
783 using ODV software (Schlitzer, 2016).

784

785 Figure 4: Principal coordinates analysis (PCoA) of all detected molecular formulae and
786 their normalized FT-ICR-MS signal intensities (averages of duplicates), based on Bray-
787 Curtis dissimilarity of all samples, colour coded by waters dominating in each layer.
788 Environmental and calculated parameters fitted to the PCoA factors are shown with
789 black bold and dashed arrows, respectively. HU = highly unsaturated compounds, UA =
790 unsaturated aliphatic compounds, Sal = Salinity, Temp = Temperature. The symbol ŧ
791 represents samples collected at 1000 m.

792

793 Figure 5: Differences of FT-ICR-MS normalized signal intensities of SPE-DOM along a
794 molecular mass scale (*m/z)* between a) the inflow of Atlantic water collected at the
795 DCM in stn 25 and the overflow of Mediterranean water represented by the LIW at stn
796 18, b) LIW collected at stn 01 (near formation site) and LIW at stn 18 in the western
797 basin. Positive differences correspond to higher relative intensities of DCM at stn 25
798 and LIW at stn 01 DOM, while negative differences correspond to higher relative
799 intensities in the LIW stn 18 DOM, respectively for a and b panels. The weighted
800 average molecular mass of these differential spectra was calculated taking into account
801 absolute peak heights of positive and negative peaks.

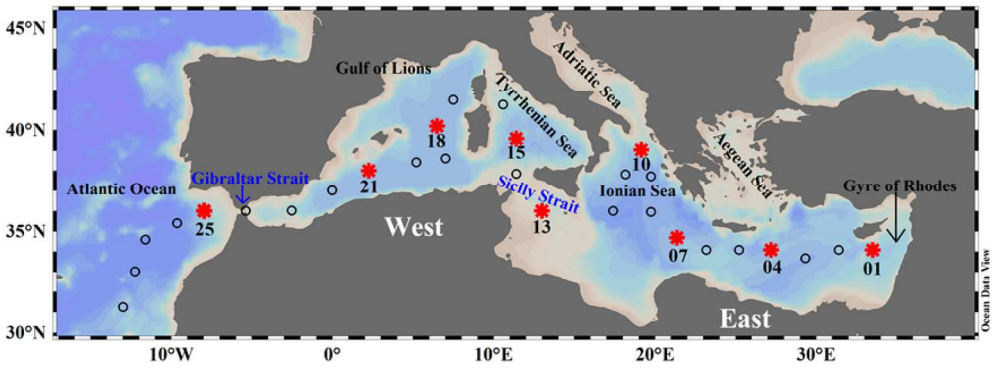
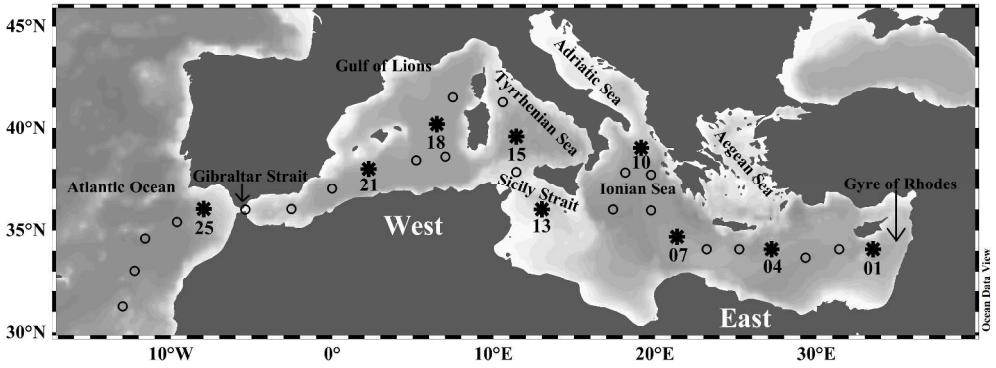


Figure 1: Study area and sampling stations. The circles depict all cruise stations and asterisks represent the stations where samples were taken for molecular characterization of the DOM. Figure created using ODV software (Schlitzer, 2016).

150x56mm (300 x 300 DPI)



Black & White version for print publication%"%"Figure 1: Study area and sampling stations. The circles depict all cruise stations and asterisks represent the stations where samples were taken for molecular characterization of the DOM. Figure created using ODV software (Schlitzer, 2016).

326x121mm (300 x 300 DPI)

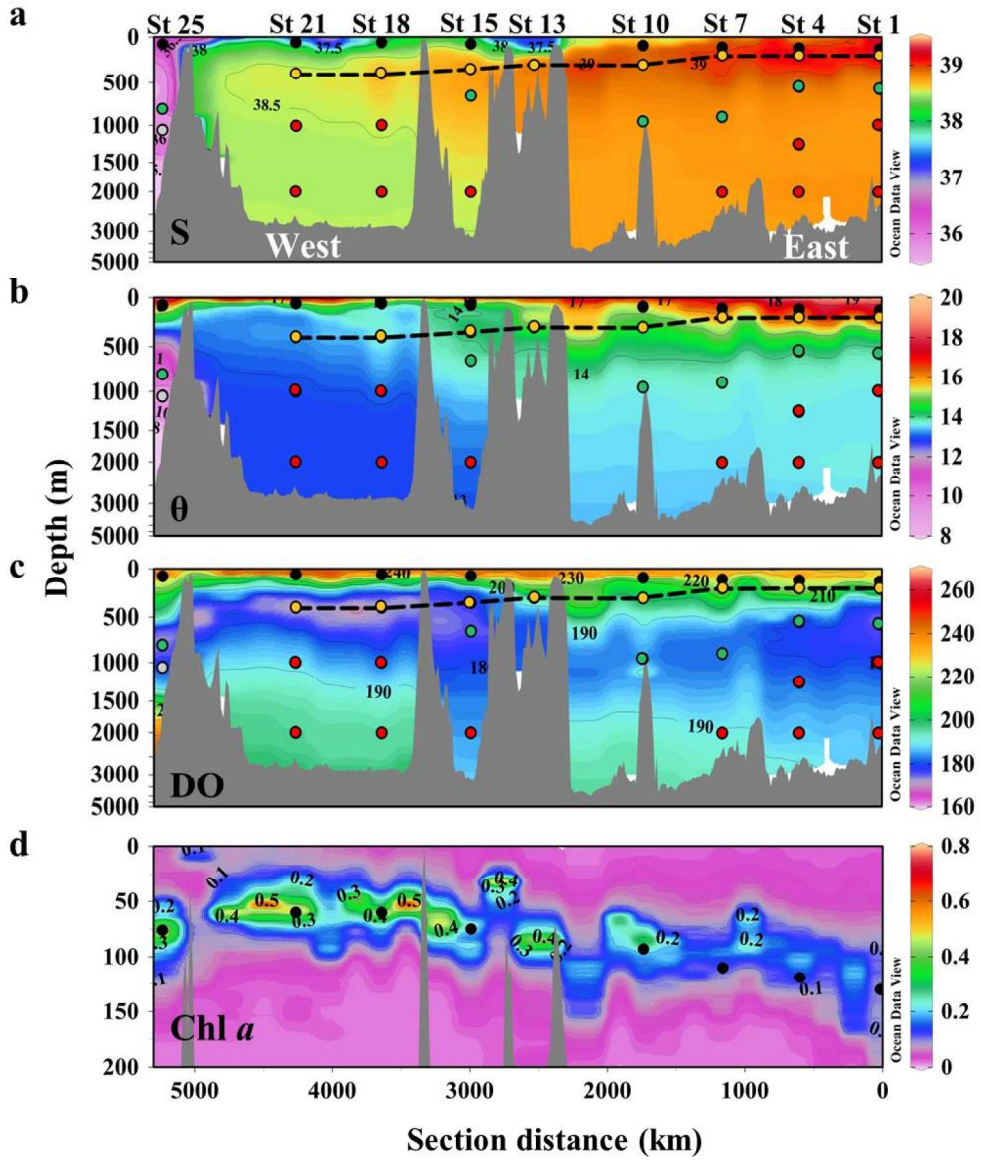


Figure 2: Distribution of (a) salinity (S), (b) potential temperature (θ) in ºC, (c) dissolved oxygen (DO) in µmol/kg and (d) fluorescence of chlorophyll a (Chl a) in mg m−3 obtained from the sensors attached to the rosette sampler along the Mediterranean Sea. Black, yellow, green and red dots represent samples taken in the epipelagic layer (DCM), Levantine Intermediate Water (LIW), oxygen minimum layer and deep waters, respectively. The dashed black line represents the route of the LIW along the transect. Note that the depth is displayed on a non-linear scale. Values from all stations were used to show these distributions. Figure created using ODV software (Schlitzer, 2016).

173x208mm (300 x 300 DPI)

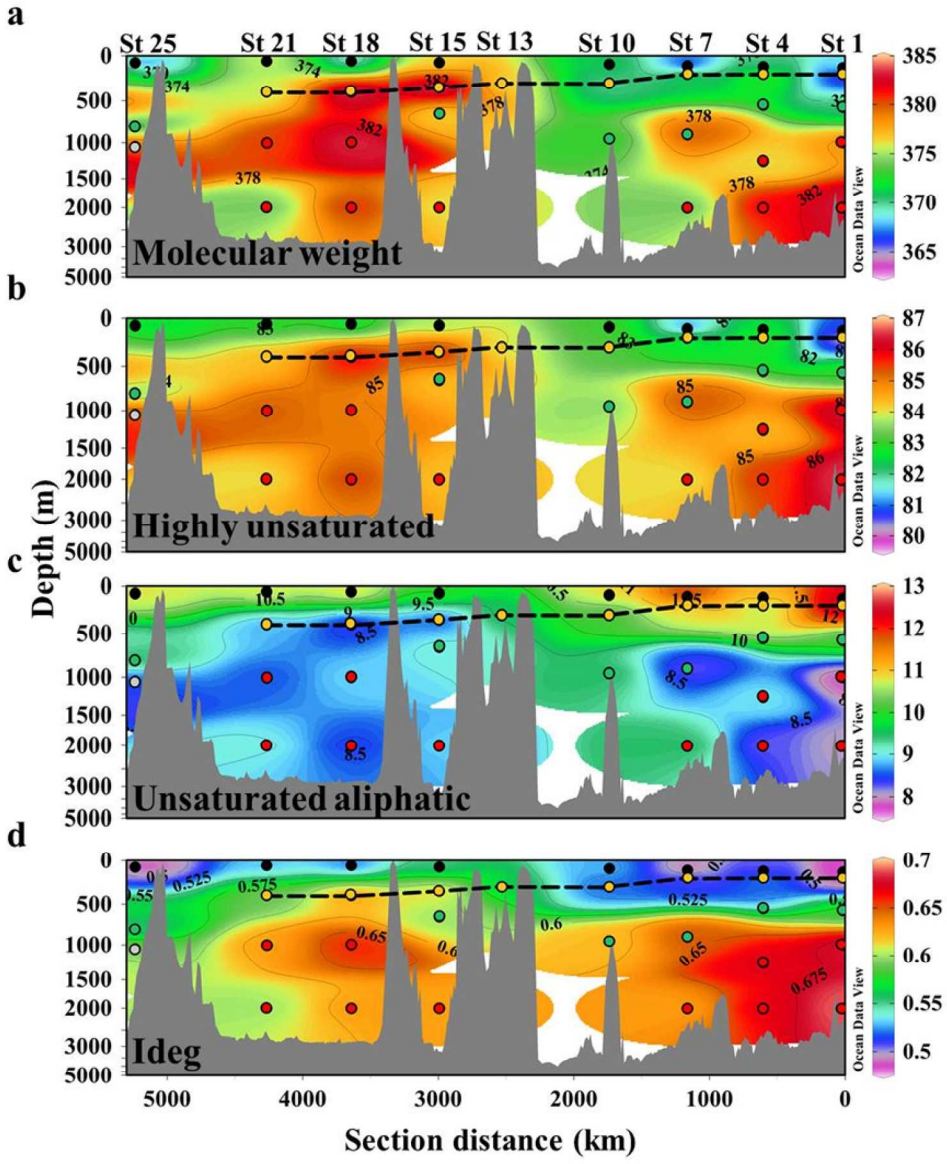


Figure 3: Distributions of (a) molecular weight in Dalton, (b) highly unsaturated compounds in %, (c) unsaturated aliphatic compounds in % and (d) degradation index (range between 0 and 1, unitless) using averages of duplicates in the Mediterranean Sea and Northeast Atlantic Ocean. The dashed black line represents the route of the LIW along the transect. Note that the depth is displayed on a non-linear scale. Figure created using ODV software (Schlitzer, 2016).

149x185mm (300 x 300 DPI)

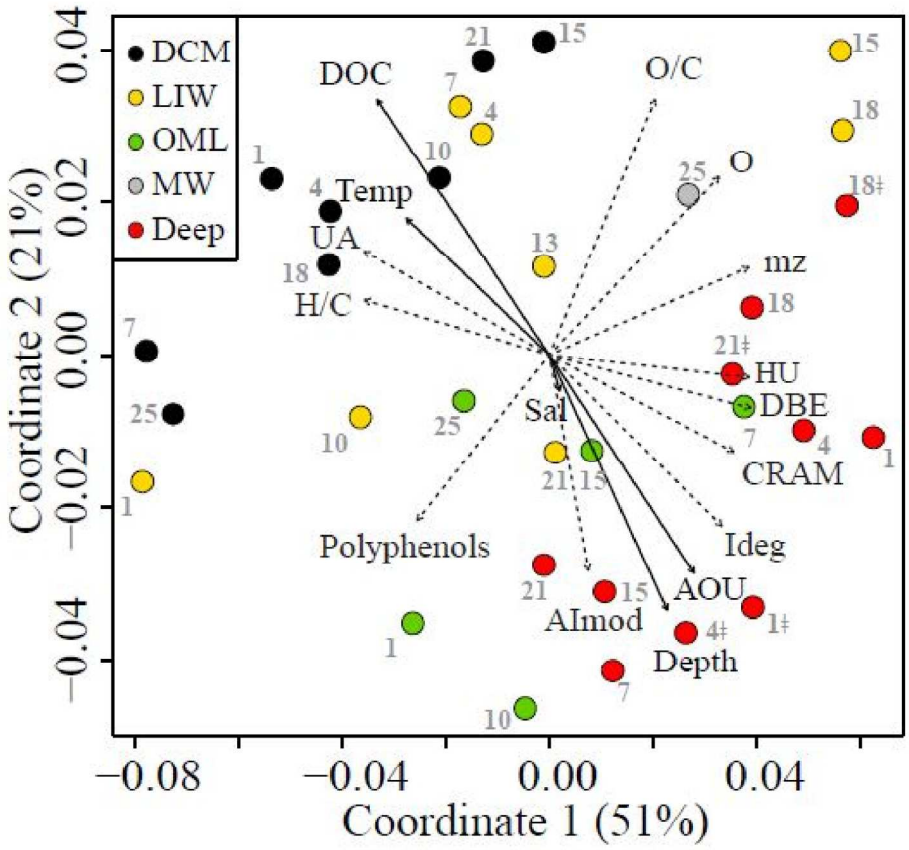


Figure 4: Principal coordinates analysis (PCoA) of all detected molecular formulae and their normalized FT-ICR-MS signal intensities (averages of duplicates), based on Bray-Curtis dissimilarity of all samples, colour coded by waters dominating in each layer. Environmental and calculated parameters fitted to the PCoA factors are shown with black bold and dashed arrows, respectively. HU = highly unsaturated compounds, UA = unsaturated aliphatic compounds, Sal = Salinity, Temp = Temperature. The symbol ǂ represents samples collected at 1000 m.

153x140mm (300 x 300 DPI)

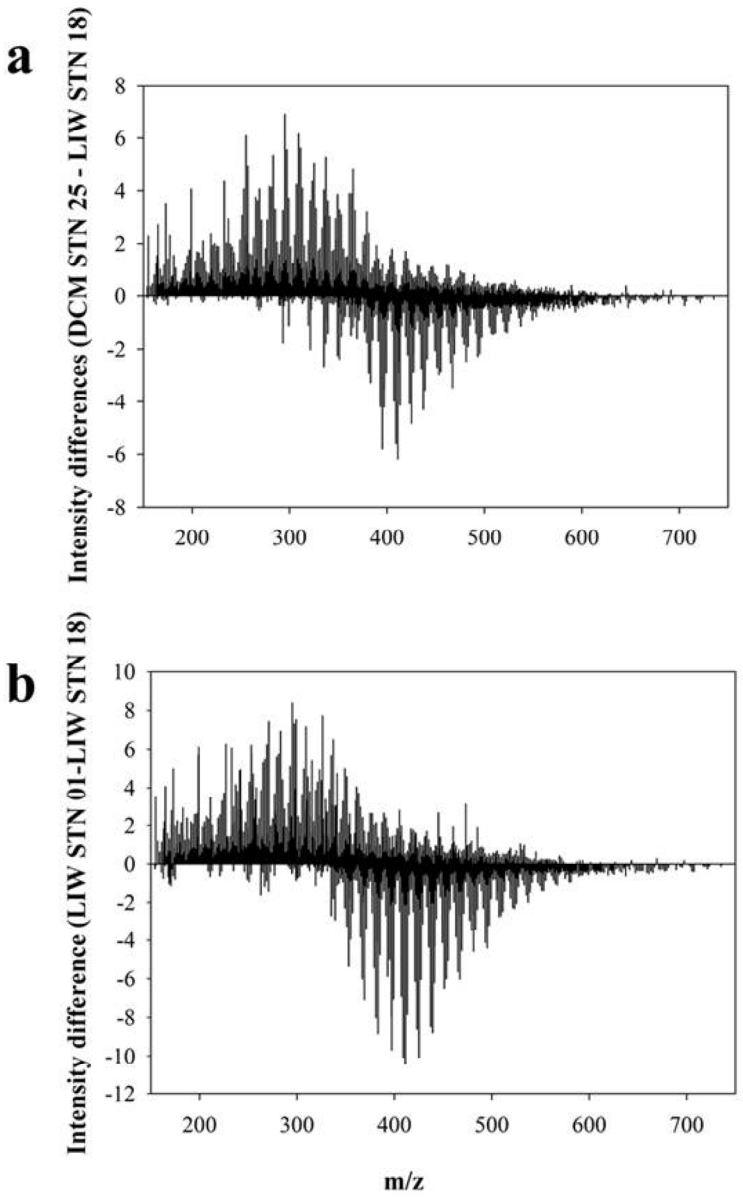


Figure 5: Differences of FT-ICR-MS normalized signal intensities of SPE-DOM along a molecular mass scale (m/z) between a) the inflow of Atlantic water collected at the DCM in stn 25 and the overflow of Mediterranean water represented by the LIW at stn 18, b) LIW collected at stn 01 (near formation site) and LIW at stn 18 in the western basin. Positive differences correspond to higher relative intensities of DCM at stn 25 and LIW at stn 01 DOM, while negative differences correspond to higher relative intensities in the LIW stn 18 DOM, respectively for a and b panels. The weighted average molecular mass of these differential spectra was calculated taking into account absolute peak heights of positive and negative peaks.

96x147mm (300 x 300 DPI)

Supplementary Table S1: Physical-chemical properties and solid-phase extraction efficiency (Ext effic) of DOM on a carbon basis. SD = standard deviation from duplicate analyses of DOC concentration.

Stn	Sample	Depth m	Sal	Pot T °C	DO µmol kg^{-1}	AOU µmol kg^{-1}	DOC µM	SD	Ext effic %
01	Deep	2000	38.78	13.62	186.7	61.3	39.7	1.1	50.5
01	Deep	1000	38.75	13.60	180.0	68.2	43.9	2.8	45.8
01	OML	575	38.80	13.82	174.4	72.6	45.7	5.4	37.3
01	LIW	201	39.17	16.87	214.3	18.1	60.2	0.9	47.7
01	DCM	130	39.22	17.63	224.9	4.0	68.8	0.9	42.2
04	Deep	2000	38.77	13.60	187.6	60.5	43.1	5.2	46.9
04	Deep	1250	38.76	13.61	183.0	65.1	43.9	2.8	48.1
04	OML	551	38.85	14.00	176.9	69.1	45.7	2.1	-
04	LIW	201	39.13	16.17	220.4	15.1	59.2	0.4	44.1
04	DCM	119	39.03	16.59	230.6	3.1	65.9	2.2	42.7
07	Deep	2001	38.75	13.51	190.6	58.0	44.9	0.8	50.7
07	OML	900	38.80	13.77	182.3	65.0	45.7	8.5	49.1
07	LIW	201	39.02	15.95	220.3	16.4	59.7	0.7	51.5
07	DCM	110	38.97	16.52	228.8	5.4	64.5	0.1	50.9
10	OML	951	38.76	13.63	184.3	63.7	42.3	1.8	51.0
10	LIW	300	38.90	14.52	201.6	38.0	58.6	0.2	44.3
10	DCM	93	38.70	15.08	230.2	10.9	60.8	6.5	43.9
13	LIW	301	38.87	14.81	196.1	46.1	54.8	0.4	46.7
15	Deep	2001	38.53	13.11	184.7	66.3	47.1	7.1	49.9
15	OML	651	38.71	13.83	174.5	72.6	49.6	6.4	44.6
15	LIW	351	38.75	14.16	176.6	68.9	46.2	1.2	51.5
15	DCM	75	38.32	14.33	242.7	2.7	64.3	5.2	41.7
18	Deep	2000	38.48	12.90	195.0	57.2	42.7	0.6	52.8
18	Deep	1000	38.53	13.11	183.0	68.1	42.7	5.4	50.1
18	LIW	401	38.67	13.77	171.3	76.2	47.4	0.7	48.4
18	DCM	60	37.96	14.23	242.8	3.7	61.0	0.2	49.2
21	Deep	2000	38.48	12.90	196.3	55.9	44.4	5.4	50.9
21	Deep	1000	38.48	12.94	189.0	63.0	42.5	5.4	51.1
21	LIW	401	38.53	13.25	169.9	80.5	47.4	0.7	48.2
21	DCM	60	37.63	14.75	243.1	1.4	65.2	1.6	42.0
25	MW	1050	35.89	9.80	181.0	93.0	45.7	0.6	51.1
25	OML	800	35.80	10.42	174.3	96.3	49.8	5.1	49.1
25	DCM	76	36.30	16.12	230.5	9.8	69.0	1.0	41.1

Supplementary Table S2: Intensity-weighted averages (average ± standard deviation) of molecular parameters from the FT-ICR-MS duplicate analysis of SPE-DOM. mw = molecular weight in Dalton, AI_{mod} = modified aromaticity index (Koch and Dittmar, 2006, 2016), DBE = double bond equivalent (Koch and Dittmar, 2006), Ideg = degradation index (Flerus et al., 2012), Poly = polyphenols, HU = highly unsaturated compounds, UA = unsaturated aliphatic compounds, CRAM = carboxyl-rich alicyclic molecules (Hertkorn et al., 2006), at = atoms. Note that sample St 7 Deep was analysed only once.

St	Sample	Depth	mw	H/C	O/C	AI_{mod}	DBE	Ideg	Poly	HU	UA	CRAM
		m	Da	at H/at C	at O/at C	-	-	-	%	%	%	%
01	Deep	2000	382.9 ± 1.3	1.296 ± 0.001	0.400 ± 0.001	0.248 ± 0.000	7.93 ± 0.03	0.685 ± 0.002	4.1 ± 0.1	86.5 ± 0.3	8.1 ± 0.1	53.88 ± 0.01
01	Deep	1000	378.3 ± 1.6	1.294 ± 0.001	0.395 ± 0.003	0.252 ± 0.001	7.90 ± 0.02	0.68 ± 0.01	4.5 ± 0.2	86.3 ± 0.5	7.8 ± 0.2	54.0 ± 0.5
01	OML	575	374.1 ± 0.6	1.307 ± 0.006	0.387 ± 0.002	0.250 ± 0.002	7.77 ± 0.01	0.64 ± 0.01	4.7 ± 0.1	84.2 ± 0.2	9.0 ± 0.1	52.5 ± 0.2
01	LIW	201	363.8 ± 0.5	1.326 ± 0.001	0.383 ± 0.001	0.244 ± 0.001	7.410 ± 0.002	0.49 ± 0.01	5.0 ± 0.1	79.81 ± 0.04	12.61 ± 0.01	49.7 ± 0.2
01	DCM	130	370.3 ± 2.4	1.316± 0.001	0.391 ± 0.003	0.243 ± 0.002	7.56 ± 0.02	0.482 ± 0.003	4.9 ± 0.4	81.2 ± 0.4	12.24 ± 0.03	50.7 ± 0.3
04	Deep	2000	381.7 ± 1.6	1.298 ± 0.000	0.396 ± 0.001	0.249 ± 0.001	7.90 ± 0.02	0.67 ± 0.01	4.6 ± 0.3	85.6 ± 0.4	8.2 ± 0.0	52.9 ± 0.1
04	Deep	1250	376.3 ± 0.9	1.302 ± 0.000	0.392 ± 0.003	0.250 ± 0.001	7.812 ± 0.003	0.665 ± 0.001	4.8 ± 0.4	84.5 ± 0.5	9.0 ± 0.2	52.4 ± 0.4
04	OML	551	-	-	-	-	-	-	-	-	-	-
04	LIW	201	374.5 ± 0.5	1.307 ± 0.001	0.397 ± 0.000	0.244 ± 0.001	7.68 ± 0.01	0.526 ± 0.001	4.8 ± 0.2	83.2 ± 0.1	10.4 ± 0.1	51.5 ± 0.3
04	DCM	119	368.9 ± 0.6	1.310 ± 0.001	0.394 ± 0.004	0.244 ± 0.000	7.57 ± 0.01	0.501 ± 0.004	4.8 ± 0.2	82.4 ± 0.6	11.3 ± 0.3	50.8 ± 0.7
07	Deep	2001	374.1	1.305	0.391	0.249	7.76	0.63	4.74	84.0	9.5	52.1
07	OML	900	379.9 ± 0.1	1.296 ± 0.000	0.396 ± 0.000	0.250 ± 0.000	7.887 ± 0.004	0.646 ± 0.002	4.82 ± 0.03	85.37 ± 0.02	8.27 ± 0.03	52.38 ± 0.02
07	LIW	201	373.6 ± 1.1	1.306 ± 0.000	0.398 ± 0.000	0.245 ± 0.000	7.67 ± 0.02	0.516 ± 0.000	4.75 ± 0.02	83.4 ± 0.1	10.33 ± 0.02	51.39 ± 0.01
07	DCM	110	363.9 ± 0.8	1.317 ± 0.001	0.385 ± 0.003	0.246 ± 0.000	7.46 ± 0.01	0.48 ± 0.01	5.3 ± 0.1	80.3 ± 0.4	12.5 ± 0.2	49.5 ± 0.6
10	OML	951	372.7 ± 1.9	1.305 ± 0.002	0.383 ± 0.013	0.252 ± 0.001	7.75 ± 0.04	0.621 ± 0.004	5.1 ± 0.1	83.6 ± 0.5	9.5 ± 0.3	51.8 ± 0.5
10	LIW	300	370.0 ± 2.4	1.305 ± 0.004	0.387 ± 0.005	0.251 ± 0.000	7.67 ± 0.05	0.52 ± 0.01	5.2 ± 0.3	83.2 ± 1.1	10.0 ± 0.5	51.6 ± 0.9
10	DCM	93	371.3 ± 0.1	1.310 ± 0.002	0.398 ± 0.001	0.243 ± 0.000	7.61 ± 0.01	0.534 ± 0.001	4.5 ± 0.2	83.1 ± 0.1	10.9 ± 0.1	51.4 ± 0.1
13	LIW	301	376.2 ± 1.5	1.306 ± 0.000	0.395 ± 0.001	0.246 ± 0.001	7.73 ± 0.02	0.555 ± 0.000	4.7 ± 0.1	83.9 ± 0.2	9.6 ± 0.1	52.0 ± 0.1
15	Deep	2001	376.1 ± 0.2	1.301 ± 0.000	0.389 ± 0.000	0.252 ± 0.000	7.81 ± 0.01	0.628 ± 0.005	5.02 ± 0.03	84.3 ± 0.1	8.9 ± 0.1	52.45 ± 0.03
15	OML	651	376.3 ± 0.8	1.301 ± 0.001	0.392 ± 0.002	0.250 ± 0.000	7.80 ± 0.02	0.60 ± 0.01	5.02 ± 0.03	84.3 ± 0.1	9.08 ± 0.04	52.2 ± 0.3
15	LIW	351	385.0 ± 3.6	1.298 ± 0.001	0.407 ± 0.004	0.244 ± 0.002	7.90 ± 0.05	0.62 ± 0.02	4.31 ± 0.4	85.9 ± 0.6	8.4 ± 0.1	52.7 ± 0.1

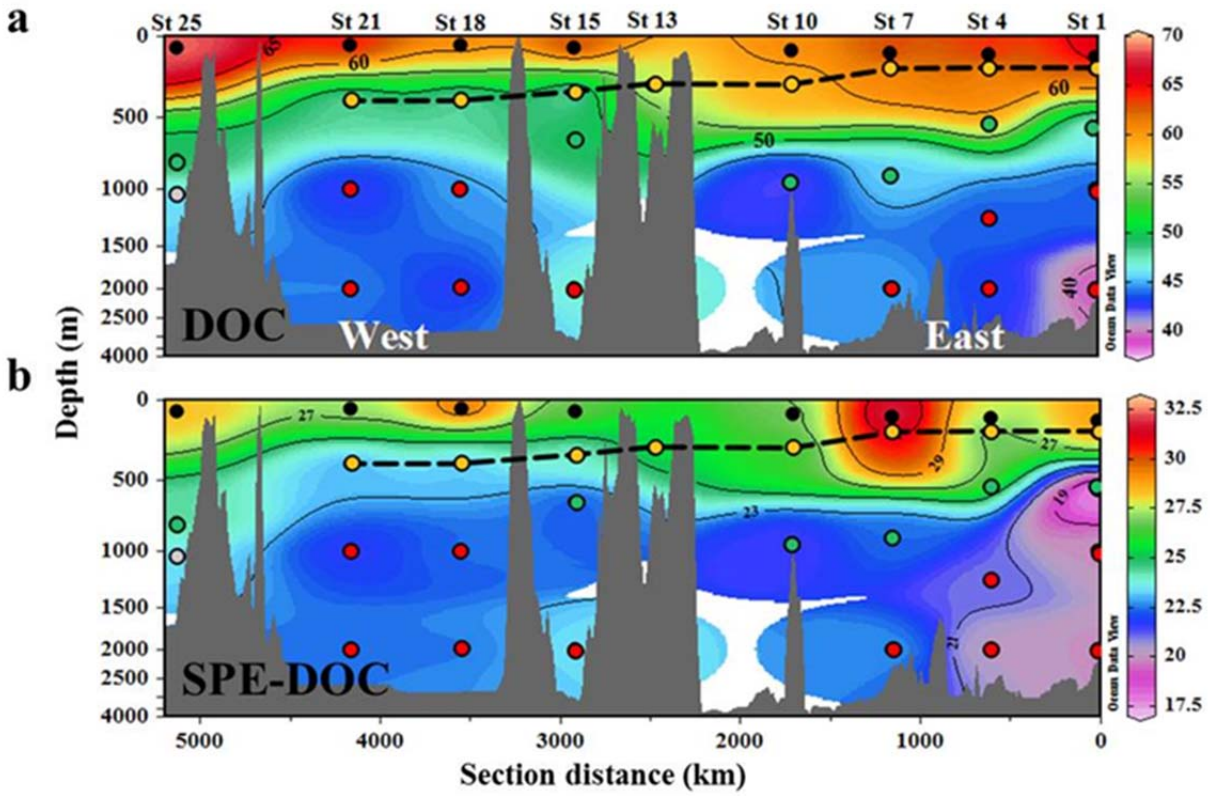
St	Sample	Depth	mw	H/C	O/C	AI_{mod}	DBE	Ideg	Poly	HU	UA	CRAM
		m	Da	at H/at C	at O/at C	-	-	-	%	%	%	%
15	DCM	75	376.3 ± 1.3	1.306 ± 0.004	0.401 ± 0.002	0.243 ± 0.000	7.70 ± 0.05	0.55 ± 0.01	4.64 ± 0.04	83.7 ± 0.6	10.1 ± 0.3	51.3 ± 0.3
18	Deep	2000	380.8 ± 0.6	1.298 ± 0.001	0.400 ± 0.002	0.248 ± 0.000	7.87 ± 0.02	0.633 ± 0.002	4.6 ± 0.1	85.5 ± 0.3	8.4 ± 0.2	52.8 ± 0.4
18	Deep	1000	383.4 ± 3.0	1.301 ± 0.002	0.402 ± 0.006	0.244 ± 0.001	7.87 ± 0.05	0.66 ± 0.01	4.5 ± 0.2	85.0 ± 0.7	8.9 ± 0.4	52.2 ± 0.7
18	LIW	401	384.1 ± 1.2	1.296 ± 0.001	0.404 ± 0.001	0.246 ± 0.000	7.92 ± 0.03	0.633 ± 0.002	4.4 ± 0.1	86.0 ± 0.2	8.28 ± 0.02	52.6 ± 0.1
18	DCM	60	368.4 ± 3.1	1.309 ± 0.005	0.394 ± 0.004	0.246 ± 0.005	7.58 ± 0.03	0.52 ± 0.02	5.1 ± 0.6	82.6 ± 0.7	10.9 ± 0.3	51.0 ± 0.3
21	Deep	2000	374.1 ± 0.9	1.302 ± 0.001	0.388 ± 0.002	0.252 ± 0.000	7.78 ± 0.01	0.602 ± 0.002	5.20 ± 0.04	84.1 ± 0.2	9.0 ± 0.1	51.7 ± 0.4
21	Deep	1000	380.4 ± 2.4	1.299 ± 0.001	0.396 ± 0.003	0.248 ± 0.001	7.87 ± 0.04	0.63 ± 0.01	4.7 ±0.2	85.3 ± 0.6	8.5 ± 0.2	52.7 ± 0.4
21	LIW	401	375.5 ± 1.1	1.302 ± 0.001	0.392 ± 0.003	0.251 ± 0.001	7.78 ± 0.01	0.59 ± 0.01	5.08 ± 0.18	84.3 ± 0.5	8.8 ± 0.3	52.0 ± 0.6
21	DCM	60	375.7 ± 3.8	1.310 ± 0.001	0.400 ± 0.006	0.242 ± 0.002	7.66 ± 0.05	0.53 ± 0.01	4.7 ± 0.3	82.9 ± 0.5	10.8 ± 0.3	51.0 ± 0.3
25	MW	1050	379.5 ± 1.9	1.297 ± 0.000	0.401 ± 0.002	0.247 ± 0.002	7.85 ± 0.02	0.59 ± 0.01	4.7 ± 0.2	85.5 ± 0.2	8.5 ± 0.1	52.4 ± 0.1
25	OML	800	372.6 ± 1.2	1.304 ± 0.000	0.389 ± 0.003	0.250 ± 0.001	7.72 ± 0.01	0.551 ± 0.001	5.2 ± 0.1	83.4 ± 0.2	9.6 ± 0.1	51.1 ± 0.2
25	DCM	76	364.9 ± 0.8	1.309 ± 0.000	0.388 ± 0.001	0.249 ± 0.001	7.56 ± 0.01	0.477 ± 0.002	5.3 ± 0.2	82.2 ± 0.4	10.87 ± 0.03	50.5 ± 0.1

Supplementary Table S3: Comparison between the average molecular composition of SPE-DOM from the Mediterranean Sea and the NEqPIW. All values are calculated from FT-ICR-MS data (average ± standard deviation). The letters represent the significant differences based on the Student's t-test for each variable ($p<0.05$). Values sharing the same letter are statistically equal. MF = molecular formula, Nr = Number.

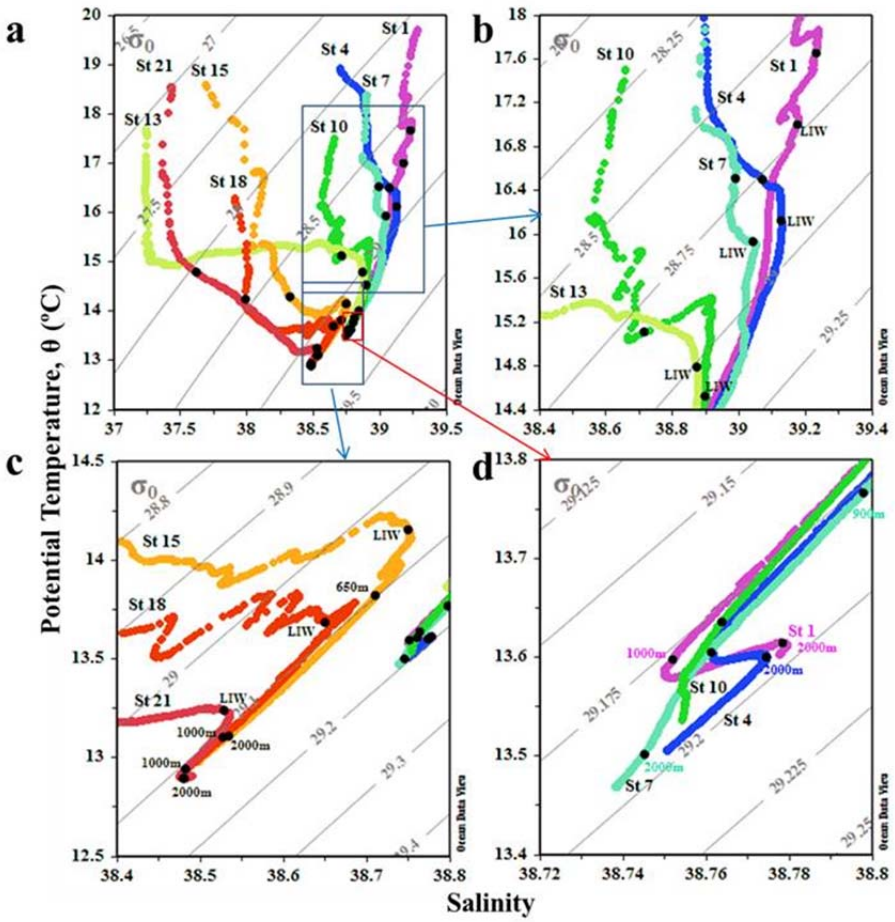
	Units	Average Mediterranean Sea n = 29	NEqPIW n = 13
General			
Number of peaks	Nr	4940 ± 338 [a]	4562 ± 112 [b]
Number of assigned MF	Nr	3105 ± 189 [a]	3142 ± 51 [a]
% of masses with MF	%	63 ± 1 [a]	69 ± 1 [b]
Weighted average mass	Da	375.3 ± 5.8 [a]	425.6 ± 1.9 [b]
Elemental composition			
C	Nr	18.82 ± 0.23 [a]	20.23 ± 0.06 [b]
H	Nr	24.52 ± 0.20 [a]	25.66 ± 0.12 [b]
O	Nr	7.40 ± 0.19 [a]	9.40 ± 0.12 [b]
N	Nr	0.36 ± 0.02 [a]	0.37 ± 0.01 [a]
S	Nr	0.07 ± 0.01 [a]	0.06 ± 0.01 [b]
P	Nr	0.01 ± 0.01 [a]	0.02 ± 0.00 [b]
H/C	at H/at C	1.30 ± 0.01 [a]	1.27 ± 0.00 [b]
O/C	at O/at C	0.39 ± 0.01 [a]	0.47 ± 0.01 [b]
C/N	at C/at N	52.39 ± 1.96 [a]	55.20 ± 2.30 [b]
Molecular indices			
AI_{mod}	Nr	0.25 ± 0.00 [a]	0.23 ± 0.00 [b]
DBE	Nr	7.74 ± 0.14 [a]	8.59 ± 0.04 [b]
Ideg	Unitless	0.59 ± 0.06 [a]	0.86 ± 0.01 [b]
Molecular diversity			
CHO	%	72.3 ± 0.6 [a]	70.8 ± 0.4 [b]
CHON	%	21.4 ± 0.7 [a]	22.3 ± 0.2 [b]
CHOS	%	5.0 ± 0.5 [a]	4.4 ± 0.6 [b]
CHOP	%	0.9 ± 0.4 [a]	2.2 ± 0.3 [b]
CHONS	%	0.2 ± 0.0 [a]	0.3 ± 0.0 [b]
CHOSP	%	0.2 ± 0.1 [a]	0.1 ± 0.0 [b]
Polyphenols	%	4.8 ± 0.3 [a]	2.6 ± 0.2 [b]
Highly unsaturated	%	83.9 ± 1.6 [a]	91.9 ± 0.5 [b]
Unsaturated aliphatics	%	9.7 ± 1.4 [a]	4.8 ± 0.3 [b]
CRAM	%	51.9 ± 1.0 [a]	48.6 ± 0.5 [b]

Supplementary Table S4: List of acronyms and definitions of water masses, molecular indices, types of molecular formulae and compounds groups.

Acronym	Description	Reference	Proxy
Hydrology			
DCM	Deep Chlorophyll Maximum		
LIW	Levantine Intermediate Water		
OML	Oxygen Minimum Layer		
WMDW	Western Mediterranean Deep Water		
EMDW	Eastern Mediterranean Deep Water		
MW	Mediterranean Water		
NEqPIW	North Equatorial Pacific Intermediate Water		
Molecular Indices			
DBE	Double Bond Equivalent	Koch and Dittmar, 2006	↑unsaturation
AI_{mod}	Aromaticity Index Modified	Koch and Dittmar, 2016	↑aromaticity
Ideg	Degradation Index	Flerus et al., 2012	↑degradation state
Molecular formula type			
CHO	Molecular formulae with C, H and O atoms		
CHON	Molecular formulae with C, H, O and N atoms		
CHOS	Molecular formulae with C, H, O and S atoms		
CHOP	Molecular formulae with C, H, O and P atoms		
CHONS	Molecular formulae with C, H, O, N and S atoms		
CHOSP	Molecular formulae with C, H, O, S and P atoms		
Compound groups			
Polyphenols	$0.5 < AI_{mod} < 0.666$		↑aromaticity
Highly unsaturated	$AI_{mod} < 0.5$, $H/C < 1.5$ and $O/C < 0.9$		Refractory DOM
Unsaturated aliphatic	$1.5 < H/C < 2$, $O/C < 0.9$ and $N = 0$		Bio-labile DOM
CRAM	Carboxyl-rich alicyclic molecules; $0.3 < DBE/C < 0.68$, $0.2 < DBE/H < 0.95$ and $0.77 < DBE/O < 1.75$	Hertkorn et al., 2006	Refractory DOM



Supplementary Figure S1: (a) DOC and (b) SPE-DOC distribution in the Mediterranean Sea and Northeast Atlantic Ocean in µmol/L. The dashed black line represents the route of the LIW along the transect. Note that the depth is displayed on a non-linear scale.



Supplementary Figure S2: θ-S diagrams of the Mediterranean stations considered for the SPE-DOM molecular characterisation. Black dots represent the 29 samples collected for this study in (a) the full range diagram and the zoom areas for (b) the intermediate waters in the eastern basin, (c) the intermediate and deep waters in the western basin and (d) the deep waters in the eastern basin.